A Comparative Assessment of Statistical Approaches for fMRI Data to Obtain Activation Maps

Research Thesis

Presented in Partial Fulfillment of the Requirements for the Graduation with Research Distinction in Statistics

by

Cristina Perez Diukina

Undergraduate Program in Statistics

The Ohio State University

2021

Thesis Committee:

Asuman Turkmen, Advisor

Oksana Chkrebtii

Abstract

Functional Magnetic Resonance Imaging (fMRI) lets us peek into the human mind and try to identify which brain areas are associated with certain tasks without the need for an invasive procedure. However, the data collected during fMRI sessions is complex; this 4 dimensional sequence of 3 dimensional volumes as images of the brain does not allow for straightforward inference. Multiple models have been developed to analyze this data and each comes with its intricacies and problems. Two of the most common ones are 2-step General Linear Model (GLM) and Independent Component Analysis (ICA). We compare these approaches empirically by fitting the models to real fMRI data using packages developed and readily available in R. The real data, obtained from an open source database *openneuro.org*, is named *BOLD5000*.

The task of interest for this thesis is image viewing versus fixation cross (resting state). We found that both the first-level GLM and ICA revealed significant activation located in the occipital lobe which is consistent with the literature on visual tasks. The second-level GLM results were consistent with the first level and found activation located in the occipital lobe as well. The Group ICA results however found activation located mainly in the temporal lobe.

Acknowledgements

Throughout the writing of this thesis I have received a lot of support, assistance, and guidance.

I would first like to thank my advisor, Professor Asuman Turkmen, whose expertise and contributions were invaluable and made this thesis possible. Secondly, I would like to thank Professor Oksana Chkrebtii for her encouragements and advice.

Table of Contents

List of Figures

Chapter 1

Introduction

This chapter presents some background on functional Magnetic Resonance Imaging (fMRI). Specifically, Section 1.1 introduces fMRI as a powerful clinical tool in the study of the brain's physical structure. This section also presents two categories of the experimental designs commonly used in fMRI studies and a brief description of the hemodynamic response function. Section 1.2 describes necessary preprocessing steps to remove sources of noise from fMRI data. Section 1.3 summarizes some of the statistical tools available for fMRI data analysis and the challenges in understanding the underlying neuronal activity associated with cognitive tasks. We also provide information about the available software for these tools. Lastly, Section 1.4 discusses the contribution of this thesis.

1.1 Functional Magnetic Resonance Imaging (fMRI)

Since its development in the early 1990s, fMRI has become the most commonly used method for the study of human brain function. fMRI is a class of imaging methods developed in order to demonstrate regional, time-varying changes that can be associated to task-induced cognitive state changes or to unregulated processes in the resting brain. It has been used in a large number of studies in the cognitive neurosciences, clinical psychiatry/psychology, and neurosurgical planning due to its

widespread availability, non-invasive nature (making it safe to use on human subjects), relatively low cost, and good spatial resolution. For an excellent introduction into the technique, we refer to the textbook by Huettel et al. [28].

The most common method of fMRI utilizes the fact that when neurons in the brain activate, the amount of blood flowing through that area increases. Thus, the activity related surplus in blood flow caused by brain activity leads to a relative increase in local blood oxygen. The signal measured in fMRI depends on this change in oxygenation and is referred to as the blood oxygenation level dependent (BOLD) response. The BOLD signal is measured from small cubic regions of the brain called voxels. Each voxel contains hundreds of thousands of neurons, so the BOLD signal measured from a voxel is indicative of the group activity of the neurons located within that voxel. As the neurons in a voxel become active due to brain function, the BOLD signal will vary over time. Therefore, the data obtained from fMRI research is a time series that is very large and complex. This data is a sequence of 3D volumes as images of the brain where the 4^{th} dimension is time. Each image usually consists of about $100,000$ voxels. During the fMRI experiment image volumes are continuously collected with a repetition time (TR) of 2–4 seconds, resulting in a total of 200–300 images for the whole time-series. The tasks are designed and timed in a manner that allows the experimenters to record measurable changes in BOLD signal in order to make inferences about task-related brain activity. The goal is to accurately characterize the BOLD signal change and to relate it to brain function.

The BOLD signal does not increase instantaneously and does not return to baseline immediately after the stimulus ends. Because these changes in blood flow are relatively slow (evolving over several seconds, usually about 6), the BOLD signal is a blurred and delayed representation of the original neural signal. The mathematical

model for the local change in BOLD response after a stimulus is presented is called the hemodynamic response function (HRF). This can be described as the ideal, noiseless response to an infinitesimally brief stimulus. It has a number of important characteristics illustrated in Figure 2.1. At the first couple seconds after the stimulus, an initial dip is observed. After a short latency of initial dip, the blood flow comes in with increasing blood volume. The increase continues gradually for about 5 seconds to reach a maximum value in Magnetic Resonance (MR) signal, called peak. After the peak, the blood flow decreases rapidly and during this period the BOLD signal falls below the baseline for a prolonged time, called undershoot. Modeling HRF is essential for the correct interpretation of neurological studies. A variety of fixed HRFs have been used including the Poisson function, the Gamma function, the Gaussian function, and double Gamma function. Standard statistical analysis of fMRI data usually calls for a "canonical" model of HRF corresponding to the double Gamma function. In this work, the canonical model is assumed for the HRF estimation since this is not the focus of the thesis.

Most fMRI studies that present a stimulus to the subjects in order to make inferences on the brain's task related activity rely on two approaches of experimental design: event-related design and block design. An event-related design presents discrete, short-duration stimuli (e.g., brief light flashes, a short sound), called events, whose timing and order may be randomized [8, 20]. Event related designs are based on the assumption that neural activity will occur for short and discrete intervals. This approach offers greater design-related flexibility to the experimenter, however, the statistical power of event related designs is inherently low, because the signal change in the BOLD fMRI signal following a single stimulus presentation is small [8]. In a block design, a condition is presented continuously for an extended time interval

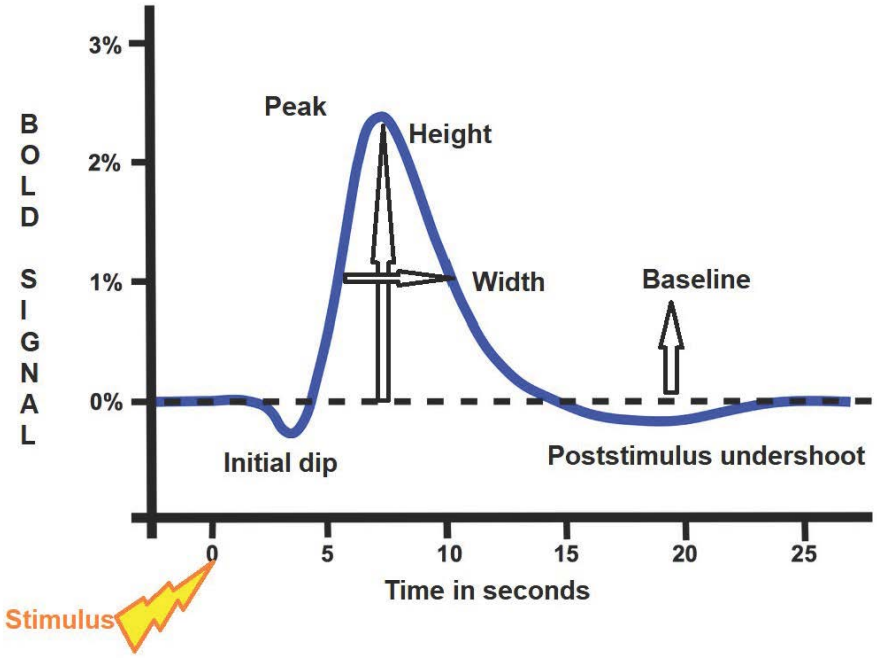

Figure 1.1: The HRF function can be described by a several characteristics including the time from the stimulus until peak, height of response, the width of the HRF at half the height, post-stimulus undershoot, and in some cases an initial dip.

(block) to maintain cognitive engagement, and different task conditions are usually alternating in time. Block designs have greater statistical power and may be more appropriate if the experimental goal is to detect subtle differences in BOLD signal across different test conditions [16, 24]. However, since the block design averages the response within the block, finer details about the time-series can be lost [44].

The experimental design is a particularly important step as it has been found to greatly influence the reliability of fMRI results [6]. Figure 1.2 presents a diagram that compares event-related and block design in a simple fictional experiment where a subject is either presented with an image (stimulus) or a fixation cross (resting state that avoids distraction). In such an experiment, we could compare the brain areas that activate when the subject is viewing an image compared to the brain areas activated when the subject is fixating a cross (our control task). Previous literature strongly suggests that an area in the occipital area should be associated with visual tasks [49]. The experiment's design complexity can of course be increased to arrive at more interesting and less documented results. Therefore, researchers also developed some optimal designs allowing one to search through the space of possible designs, with the dimensionality of the space defined by the number of design parameters allowed to vary [39]. Although optimal designs of fMRI experiments is helpful to increase the signal-to-noise ratio, using such designs is generally more computationally challenging compared to two formerly described designs.

fMRI is unquestionably a powerful tool to detect functional activation within the brain, but the obtained data from fMRI experiences cannot be easily analyzed. The signal is relatively weak and various sources of noise in the data must be carefully controlled. These sources include thermal noise, power fluctuations, variation in subject cognition, head motion effects, physiological noise (induced by respiration

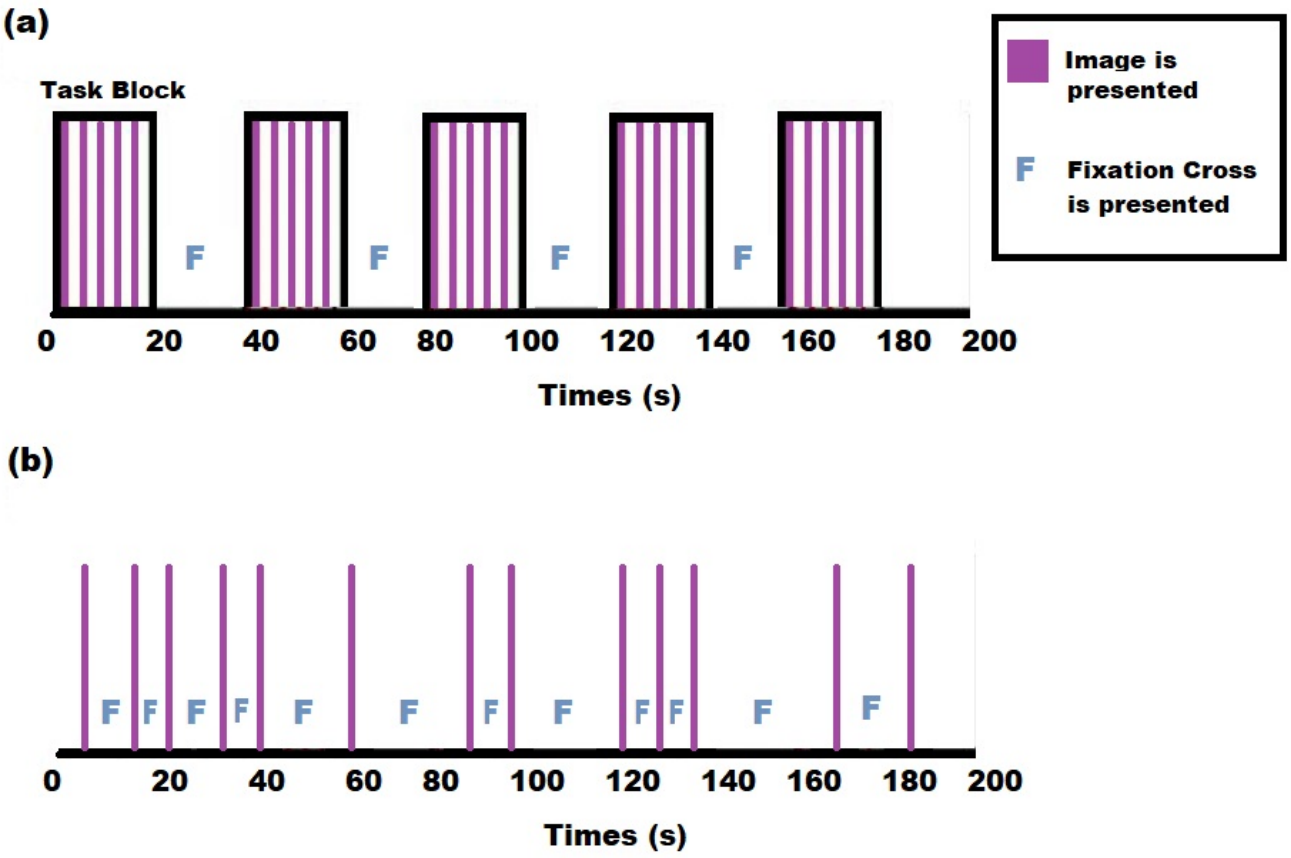

Figure 1.2: Diagrams showing a representation of a simple fictional experiment where a subject is either presented an image (visual task) or a fixation cross (resting state): (a) Blocked design with 5 task blocks where images are presented and 4 fixation (control) blocks, (b) Event-related design with jittered inter-stimulus-interval. The purple lines indicate when the images are presented to the subject.

and heart-beat) and artifact-induced problems. Furthermore, the brain is constantly at work so it is not directly observable which zones of activation are related to a certain experiment. Because of these factors, preprocessing steps and powerful data analysis must be applied to find task related activations within the brain. These concepts will be discussed further in the following subsections.

1.2 Preprocessing

fMRI is a commonly used technique for mapping human brain activity. However, the BOLD response induced by neuronal activity only represents a relatively small percentage of the variance of the signal. In general, the measured BOLD signal is very small compared to the total intensity of the actual signal, and compared to the total spatial and temporal variability across scans. Therefore, there is a need to remove the noise sources via preprocessing before the statistical analysis [55] in order to reduce effect of noise on the data, and to retrieve the rather small BOLD signal component. Preprocessing affects the data and how it can be analyzed, there is, however, no consensus on how it does so and what is the best series of steps [12, 13, 61]. Most pipelines include slice-time correction, head-motion correction, co-registration and normalization, as well as spatial smoothing.

Slice timing correction accounts for the fact that slices, composing the total volume of the brain, are obtained at different times, and as a result, are temporally misaligned from each other. Slice timing correction (STC) is the preprocessing step applied to correct for these slice-dependent delays, achieved by shifting the time series of each slice to temporally align all slices to a reference time-point. Slice timing correction is particularly important when the times to repetition (TRs) are long and

the expected hemodynamic response may vary significantly between slices. Although the effectiveness of STC can interact with other preprocessing steps in the pipeline and scanning parameters, it remains a necessary step to correct temporal misalignment [47]. Another important step is the correction of head motion. Even at the shortest TRs and when scanning the most cooperative participants, the position of the head with respect to the scanner will change to some degree. Motion causes a spatial misalignment in the source of the BOLD signal measured in all voxels at a certain TR and its neighbors in the time series. This preprocessing step is performed to make sure that each voxel represents a unique part of the brain.

Other common steps in data preprocessing include co-registration, normalization and smoothing of obtained images. Co-registration aims to have better resolution yielding the identification of the activations in the subject's individual brain by aligning the functional images with anatomical images that have greater spatial resolution. Normalization is the process of mapping the obtained image into a normalized anatomical space. It allows one to generalize the results to a larger population and to make comparisons between other studies and subjects while increasing statistical power. Finally, smoothing the data helps to improve signal-to-noise ratio and to make data close to normal so that statistical analysis requiring normality assumption can safely be applied on the smoothed data. There are some controversial opinions on the use of smoothing due to arbitrary choice of smoothing filter. As discussed in Lazar [34], one common approach is to compare the results with and without smoothing in order to understand its influence on the fMRI analysis results.

1.3 Analyzing fMRI Data

The data obtained from fMRI studies are of a highly complex nature, displaying both spatial and temporal correlation, as well as high levels of noise from varying sources, and consequently, the statistical analysis of fMRI data poses many challenges. The magnitude and complexity of the data make it difficult to create a full statistical model for describing its behavior in terms of computational feasibility and efficiency [34, 38]. The detailed discussion of the problems including but not limited to design issues, size and collection of data, potentially high correlation among observations, large amount of noise relative to signal and preprocessing can be found in Lazar et al. [35].

Statisticians have been playing an important role in fMRI studies by helping to design experiments, by improving image reconstruction techniques, by finding more efficient ways of dealing with the noise in the images, either via models or through improved computational methods and of course, by coming up with more sophisticated tools for data analysis. Since the rise in popularity of fMRI studies many statistical techniques have been developed to study brain activation. The analysis of fMRI data through the application of suitable statistical methods aims to localize regions of the brain activated by a task, to identify networks that correspond to brain function (connectivity) and to make predictions/classifications about psychological or disease states.

Activation studies focus on characterizing the neural responses to experimental tasks, which may be visualized as maps of distributed patterns of brain activity. Other common objectives in activation studies are to detect differences in patterns of brain activity among various experimental stimuli, among different subgroups of subjects,

and between two or more sessions. While the main focus of activation studies is localizing brain activation, functional connectivity studies are also popular and seek to determine multiple brain networks that show similar temporal task-related activity profiles.

The statistical approaches for activation studies can be divided into two classes: model based approaches like the general linear model (GLM) [22, 38, 53], and model free approaches like blind source separation such as Independent Component Analysis (ICA) [29]. These two approaches will be described thoroughly in Section 2. Functional connectivity, defined as the temporal dependency of neuronal activation patterns for anatomically separated brain regions, reflects statistical dependencies between distinct and distant regions of information processing neuronal populations. Therefore, it is simply a statistical concept which relies on statistical measures such as correlation, covariance, or spectral coherence. For functional brain connectivity studies, two broad classes may be identified, namely knowledge-based (or supervised) and data-driven (or unsupervised) methods which can be subdivided further into decomposition methods and clustering techniques [37]. Since investigating differences in fMRI data between cases and controls for disorders such as autism may provide new insights into disease mechanisms, several approaches commonly encountered in the machine learning literature have been proposed for brain fMRI data, such as k-nearest neighbors [57], Fisher linear discriminant [17], linear support vector machines [14], Gaussian support vector machines [23], Adaboost [40], random forests [1], and neural networks [2].

There are variety of software packages that are able to perform all aspects of analysis of an fMRI study [50]. SPM (Statistical Parametric Mapping) is an open source Matlab script software for fMRI analysis developed by Karl Friston and colleagues

[25]. FSL (FMRIB Software Library) has gained popularity in recent years due to its implementation of a number of cutting-edge methods, visualization tools and its ability of integration with grid computing yielding computational speed [54] . AFNI (Analysis of Functional NeuroImages) is another popular software that has very powerful and flexible visualization abilities while its statistical modeling and inference tools have historically been less sophisticated than those available in SPM and FSL [15]. In addition to these softwares, R and Matlab are two other cross-platform, high level scripting languages that have been used for fMRI data analysis. In this thesis, our main focus will be activation studies and their implementations available in R [56] that is an open source and a free platform.

1.4 Organization of the Thesis

The rest of the thesis is structured in three chapters. In Chapter 2, two commonly used statistical methods in fMRI activation studies, GLM and ICA, are described and a comparison based on existent literature is provided. In the next part represented by Chapter 3, BOLD 5000 data, results of GLM and ICA applied on BOLD 5000, and empirical comparison of these results are given. Finally, in Chapter 4 we provide a conclusion.

Chapter 2

Statistical Analysis

The statistical analysis of fMRI data involves several challenges since it is huge, noisy data displaying a complicated spatial and temporal structure. Therefore, statistics plays a significant role in functional neuroimaging research and in its interplay with other fields, such as neuroscience and imaging physics. Developing statistical methods based on the fMRI process modeling opens the door to more accurate analysis methods, and consequently attributing accurate scientific interpretations to results ensuring the reliability of fMRI studies.

In general, there are two common functional neuroimaging research goals: detecting brain regions that reveal task-related alterations in measured brain activity (localizing brain activation) and identifying highly correlated brain regions that exhibit similar patterns of activity over time (brain connectivity) [38]. In this chapter, we underline two popular statistical procedures for analyzing fMRI data to detect localized brain activations.

Since the development of fMRI, a variety of univariate (separated for each voxel) and multivariate methods for analyzing fMRI data have been developed to localize regions of the brain activated by specific tasks. Some of the most popular methods are the univariate, model based approaches that rely on the General Linear Model (GLM) [22, 38, 53] and the multivariate, model-free approach based on Independent Component Analysis (ICA) [29]. For the rest of this chapter, GLM and ICA methods

for detecting activated voxels in fMRI studies are described in detail. We provide a literature review on the comparison of these two approaches.

2.1 General Linear Model (GLM)

The General Linear Model (GLM) is the most widely used parametric approach to separate noise from stimulus induced signal in fMRI studies. This approach was first introduced by Friston et al. in 1994 [22]. The name "general" suggests, this model can be used for many different types of analyses, such as one-sample t -tests, two-sample t -tests, and analysis of variance (ANOVA). More formally, the GLM is a hierarchical parametrical model. When using this model for fMRI data analysis we usually fit two models: Within-Subject (Single-Subject) and Between-Subject models. This is often done in stages where the Within-Subject model is considered to be the 1st level of the analysis and the Between-Subjects model is the 2nd level. The hierarchical model then combines both stages into a single model in order to do inference and localize regions activated by a task of interest. The details of both stages are given in the subsections below. The GLM method for analyzing fMRI data is readily available in popular toolboxes such as FSL and SPM [3, 32, 48] allowing experimental scientists to analyze the data with relative ease. In this work, "*fmri.lm*" function in the "*fmri*" [51] package in R are used. The description of the GLM is given based on the algorithm implementation in the *fmri* package.

2.1.1 The Model

Typically, fMRI data can be modeled as a sum of responses, drift and noise. In this study, we follow the notation introduced in Lindquist (2008) [38]. Here the response

(i.e., BOLD response) is a linear combination of responses from K different stimuli. That is, the response for voxel i, $i = 1, 2, \ldots, V$, at time t, $t = 1, 2, \ldots, T$, for the subject j, $j = 1, 2, \ldots, M$, can be written as

$$y_{ij(t)} = \sum_{g=1}^{G} z_{ijg}(t)\gamma_{ijg} + \sum_{k=1}^{K} x_{ijk}(t)\beta_{ijk} + \epsilon_{ij}(t). \tag{2.1}$$

In equation 2.1, the first term corresponds to the drift component that accounts for drifts over time due to systematic effects. The drift can be linear or nonlinear, hence a flexible polynomial model is often employed to allow for nonlinear effects in the drift. It is very common to consider $z_{ijg}(t) = \gamma_{ijg}t^{g-1}$, i.e. p^{th} order polynomial function where $G = p$ in the model 2.1. The sum of response component of the model in 2.1 is the second term of the model which can be rewritten as

$$\sum_{k=1}^{K} x_{ijk}(t)\beta_{ijk} = \sum_{k=1}^{K} \beta_{ijk} \int h_{ij}(u)v_k(t-u)du \tag{2.2}$$

where h_{ij} is the HRF and $v_k(t)$ is the stimulus function depending on the experimental stimulus that is described by a task indicator function. We assume that the HRF function is known and defined as the difference of two gamma functions (canonical) given as

$$h(t) = (\frac{t}{d_1})^{a_1} exp(-\frac{t-d_1}{b_1}) - c(\frac{t}{d_2})^{a_2} exp(-\frac{t-d_2}{b_2}) \tag{2.3}$$

where $a_1 = 6$, $a_2 = 12$, $b_1 = 0.9$, $b_2 = 0.9$, and $d_i = a_i b_i$ (i = 1, 2), and $c = 0.35$. ϵ_{ij} is assumed to follow an AR(1), a linear model that predicts the present value of a time series using the immediately prior value in time. Specifically, $\epsilon_{ij}(t) = \rho_{ij}\epsilon_{ij}(t-1) + \psi_{ij}$ where $|\rho_{ij}| < 1$ and ψ_{ij} are independent and identically distributed errors.

2.1.2 First level: Within-Subject Model

Assuming that the drift component is accounted for, we can drop the term from the model in 2.1 and it can be rewritten in a matrix form for subject j at a given voxel as:

$$y_j = X_j\beta_j + \epsilon_j \tag{2.4}$$

where y_j is the $T \times 1$ vector containing the BOLD time series, X_j is the $T \times K$ design matrix columns corresponding to the predicted BOLD response for each condition, β_j is a $K \times 1$ vector of parameters, and $\epsilon_j \sim N(0, \sigma_j^2 V_j)$ where the covariance matrix $V_j = (v_{j(mn)}) = \rho_j^{|m-n|}$ for $|\rho_j| < 1$ corresponding to an AR(1) process which is considered as sufficient for fMRI experiments [60] for j=1,2,...,M, $m \leq T$ and $n \leq T$. The model for an imagined voxel is illustrated in Figure 2.1.

The autocorrelation coefficients (ρ_j's) are estimated from the residual vector $r_j = (r_{j1}, \ldots, r_{jT})$ of the fitted model 2.4 as

$$\hat{\rho_j} = \frac{\sum_{t=2}^{T} r_{jt} r_{j(t-1)}}{\sum_{t=1}^{T} {r_{jt}}^2}. \tag{2.5}$$

After applying a bias-correction procedure described in [58], prewhitening is used to transform model 2.4 into a linear model with approximately uncorrelated errors. The prewhitened linear model is obtained by multiplying the terms in model 2.4 with $\tilde{V_j}$ using bias corrected estimate of ρ_j and, finally, least squares estimates of the β_j's are obtained from the prewhitened model as:

$$\hat{\beta_j} = ({\tilde{X_j}}' \tilde{X_j})^{-1} {\tilde{X_j}}' \tilde{y_j} \tag{2.6}$$

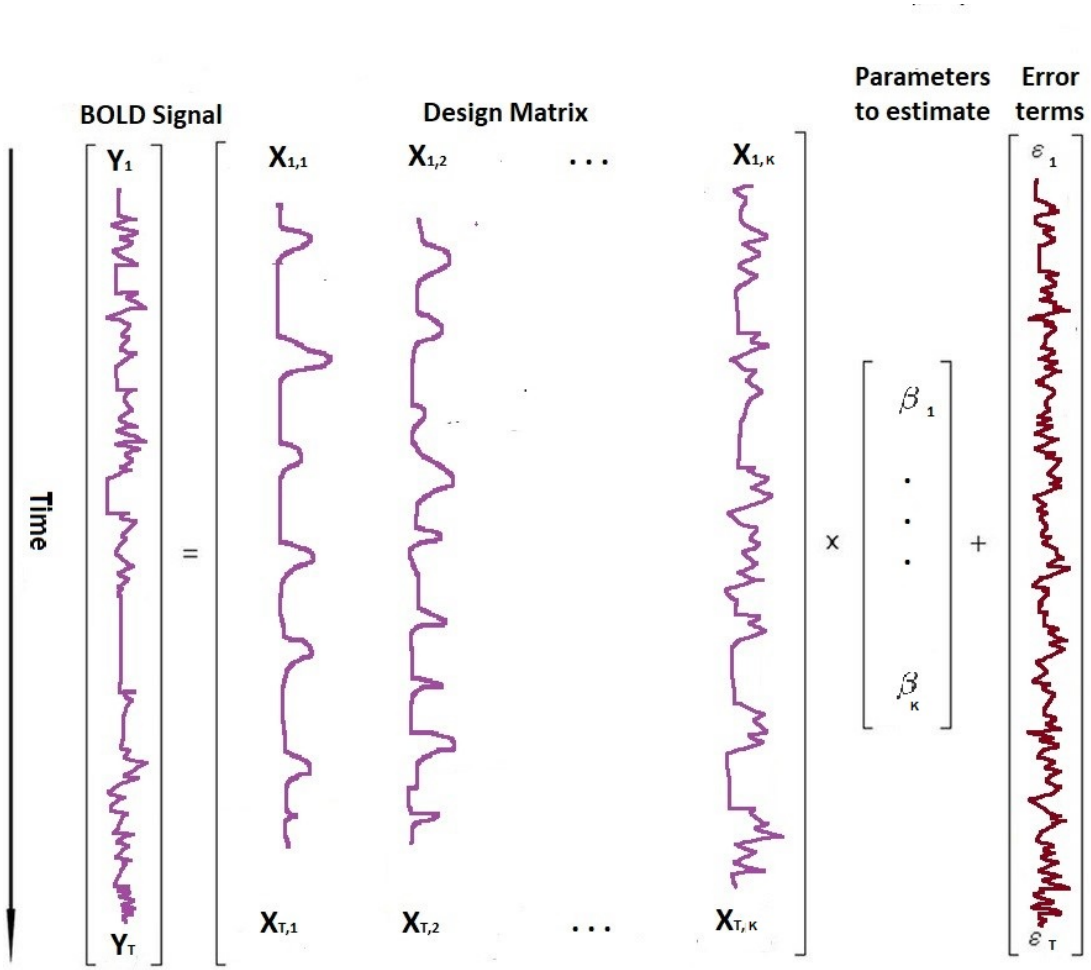

Figure 2.1: Illustration of a GLM model for one imagined voxel.

where $\tilde{X}_j = \tilde{V}_j^{-1/2} X_j$ and $\tilde{y}_j = \tilde{V}_j^{-1/2} y_j$ for $j = 1, 2, \ldots, M$. For considering more than one stimulus and to estimate a linear combination of coefficients (contrast) $\gamma_j = c'\beta_j$, one can use $\hat{\gamma_j} c' \hat{\beta}_j$ where c is a $K \times 1$ vector of contrasts for $j = 1, 2, \ldots, K$.

In summary, subject-specific regression models are fit at each voxel separately in the first level. Coefficients or contrasts are then estimated from the fit. For each subject, as there is one contrast per voxel (one coefficient per voxel), the resulting collection of voxel-specific contrast values is referred to as the subject specific contrast image, and is usually stored as a 3D image.

2.1.3 Second level: Between-Subject

After single-subject data has been analyzed for a set of participants, individual results for a given voxel are aggregated to assess commonality and stability of effects within or across groups of interest [58]. Ideally, one would regress the parameter of interest γ on a group model:

$$\gamma = X^* \beta^* + \epsilon^* \tag{2.7}$$

where X^* is $M \times q$ group-level design matrix, β^* is the group level parameter vector and ϵ^* is the group error vector with $\text{Var}(\epsilon^*) = {\sigma_*}^2 I_M$, where ${\sigma_*}^2$ is the between-subject variance and I_M is the $M \times M$ identity matrix. While X^* is often just a column of ones (for a one-sample t-test) it can take any form in general as can be seen in Figure 2.2 which is borrowed from [50].

However, since γ is unknown, we replace it with $\hat{\gamma} = (\hat{\gamma_1}, \hat{\gamma_2}, \ldots, \hat{\gamma_M})$. Therefore the model becomes:

$$\hat{\gamma} = X^* \beta^* + (\hat{\gamma} - \gamma) + \epsilon^* = X^* \beta^* + \epsilon_\gamma, \tag{2.8}$$

Test Description	Order of data	$X^*\beta^*$	Hypothesis Test
One-sample t-test. 6 observations	G_1 G_2 G_3 G_4 G_5 G_6	$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} \beta_1 \end{pmatrix}$	H_0: Overall mean=0 $H_0: \beta_1 = 0$ $H_0: c\beta = 0$ $c = [1]$
Two-sample t-test. 5 subjects in group 1 (G1) and 5 subjects in group 2 (G2)	$G1_1$ $G1_2$ $G1_3$ $G1_4$ $G1_5$ $G2_1$ $G2_2$ $G2_3$ $G2_4$ $G2_5$	$\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_{G1} \\ \beta_{G2} \end{pmatrix}$	H_0: mean of G1 different from G2 $H_0: \beta_{G1} - \beta_{G2} = 0$ $H_0: c\beta = 0$ $c = [1 \ -1]$
Paired t-test. 5 paired measures of A and B.	A_{S1} B_{S1} A_{S2} B_{S2} A_{S3} B_{S3} A_{S4} B_{S4} A_{S5} B_{S5}	$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_{diff} \\ \beta_{S1} \\ \beta_{S2} \\ \beta_{S3} \\ \beta_{S4} \\ \beta_{S5} \end{pmatrix}$	H_0: A is different from B $H_0: \beta_{diff} = 0$ $H_0: c\beta = 0$ $c = [1 \ 0 \ 0 \ 0 \ 0 \ 0]$

Figure 2.2: Examples of GLM models for particular study designs including: One-sample t-test, two-sample t-test, and paired t-test [50].

where ϵ_γ is the mixed-effects error, containing variation from both imperfect intra-subject fit, i.e. $(\hat{\gamma} - \gamma)$, and the distribution of true responses in the population, i.e. ϵ^* with $Var(\epsilon_\gamma) = Var(\hat{\gamma}) + {\sigma_*}^2 I_M$. After using this mixed-effect model, one can estimate group parameter β^* and test statistically whether it is null or not. A voxel is assigned the label "active" for a contrast of stimuli, if the estimated parameter β^* significantly deviates from zero.

One challenge present in the analysis of functional neuroimaging data is that tests of hypotheses are conducted at the voxel level, often resulting in hundreds of thousands of tests. Consequently, additional measures must be taken to maintain a reasonable type-I error rate, since it may become inflated due to the large number of tests performed. For instance, if there are 25,000 voxels to test (i.e., V=25,000), at a significance level of $\alpha = 0.05$ for the voxelwise test means that the expected value of active voxels is 1,250, even if the null hypothesis is true everywhere. This problem is known as the multiple comparison problem, and there are several strategies to tackle this problem [36]. Bonferroni correction is a multiple-comparison correction used when several independent statistical tests are being performed simultaneously [27]. Since it is not realistic to assume tests are independent given spatial location of the voxels, there are spatially dependent tests based on spatial smoothing methods [59]. The *fmri* package in R has a function called "*fmri.pvalue*" that implements both approaches.

2.2 Independent Component Analysis (ICA)

An alternative approach for an fMRI analysis is based on the independent component analysis (ICA) [29], a statistical method that aims to decompose a complex multi-

variate signal into simpler and independent subcomponents. Popular approaches for performing ICA such as maximization of information (i.e., maximum likelihood estimation), maximization of non-Gaussianity, and minimization of mutual information, are all optimization based. The most commonly used ICA algorithms are Infomax [5], FastICA [30], and joint approximate diagonalization of eigenmatrices (JADE) [10]. Since these algorithms typically work well for symmetric distributions and are less accurate for skewed distributions, some nonparametric and kernel alternatives of ICA have been developed as well [4, 7]. All these developments in ICA have resulted in a number of practical applications in biomedical problems, text document analysis, sensor signal processing, and image processing.

In the fMRI setting, ICA is used to understand the spatio-temporal structure of the signal, and it can be used to discover either spatially or temporally independent components. To identify a number of unknown sources of signal, ICA assumes that these sources are mutually and statistically independent in space or time. Therefore, ICA can be applied to fMRI data in two different ways: spatial ICA (sPCA) or temporal ICA (tICA). The first application of ICA to fMRI data used sICA [42, 43] which searches for components that are maximally independent in space. For fMRI data set analyses, sICA is generally preferred because the number of time points is small compared to the number of voxels (spatial points) making tICA typically much more computationally demanding than sICA for fMRI applications.

As in GLM analysis, ICA is available in packages such as FSL and SPM [32]. In this work, the "*fmri.sICA*" function in the "*fmri*" [51] package in R is used. The description of the ICA is given based on the algorithm implementation in the *fmri* package.

2.2.1 The Model

Let Y_j be an $T \times V$ matrix containing the centered BOLD values with rows of zero mean for the individual j for $j = 1, 2, \ldots, M$. The purpose of the analysis is to factor the data matrix, Y_j, into a product of a set of time courses and a set of spatial patterns. Therefore, for the observed Y_j, we want to estimate a $T \times R$ matrix A (mixing matrix) and $R \times V$ matrix S (spatially independent components) so that

$$Y_j = A_j S_j.$$

Thus, the ICA decomposition of Y_j can be defined as an invertible transformation

$$S_j = W_j Y_j$$

where W_j is pseudo inverse of A_j and is called "unmixing" matrix.

In typical ICA analysis, it is generally assumed that the number of sources is equal to number of mixtures which is T in case of sICA [31]. To remedy this problem, a PCA based data pre-processing is generally used where the unmixing matrix becomes a square matrix of size $T \times T$ and this process is called "prewhitening" the data. In real world applications, it is typical to apply ICA after a preliminary dimension reduction of the input data-matrix.

In this study, we use FastICA algorithm [30] that searches for (maximally) non Gaussian sources, where non-gaussianity of the extracted sources is maximized. To measure non-Gaussianity, FastICA relies on nonlinear functions such as $f(u) = logcosh(u)$ or $f(u) = e^{\frac{-u^2}{2}}$.

2.2.2 Group ICA

The analysis described in the previous section gives results for one subject. Several ICA multi-subject analysis approaches have been proposed in the literature. In this thesis, we use the method proposed by Esposito et al. (2005)[21]. This approach performs single-subject ICA for each subject (or run) and then combines the output into a group using self-organized clustering.

Specifically, the ICA estimates from each subject were organized in one single set of components (i.e., S and W) with an additional label preserving the link from the components to the original subject; those components were then clustered according to their mutual similarities. A natural measure of similarity between the estimated independent components is the absolute value of their mutual correlation coefficients that can be defined by a weighted sum of the the spatial and temporal correlations of the components. Here, spatial correlations are calculated using the columns of the components (S_j) whereas temporal correlations are calculated using the columns of the mixing matrix (A_j). The weight parameter is bounded between 0 and 1 and allows a user defined weighting of temporal or spatial similarity of the components. After calculating similarity matrices for each subject, these matrices are transformed into dissimilarity matrices of size $T \times T$. Then a supervised hierarchical clustering algorithm, linking the components to each other only when differently labeled (i.e., belonging to different subjects), is implemented.

Once the estimates belonging to a cluster have been retrieved, the average component of this cluster is computed and, henceforth, assumed as the group component representative of the cluster. Then the group components can be combined using the group proportions as weights. The described group ICA is implemented in

"*fmri.sgroupICA*" function in the *fmri* package in which only spatial similarity of the components is used.

A method for performing comparisons of group ICA data is proposed in [9]. After constructing ICA components, one might select a "component of interest" by choosing the component that correlates the highest with a task waveform. Next, a test is performed to determine which voxels are significantly "contained" in this component and these are identified to be "activated". In this study, the "ICAfingerprint" function in the fmri package is used to select a component which is then used to identify activated components. This function implements a method proposed by Martino et al. [19].

2.3 GLM versus ICA

Most GLM models fit to fMRI data make several assumptions that are often viewed as unrealistic in practice but must be met for inferences to be valid. First, the voxels are assumed to independent. It is, however, reasonable to assume that there is some degree of spatial correlation, in other words that voxels in close neighborhoods are more likely to be activated at the same time. It is important to note that there are more voxels than time points and fitting a truly multivariate model leads to issues with parameter estimation. When using whitening, or decorrelation, it is assumed that the true error correlation is known, whereas in practice it is usually estimated from the data. This estimate can be biased and highly variable. Finally, the same model is fit to every voxel. The parameters that are estimated will differ since the models are fit independently at every voxel. It can be argued that different locations would necessitate distinct models to better capture the complexity of the brain. In section

11.3, Nicole Lazar describes approaches for model selection in order to fit the most appropriate model at different locations [34]. There are also assumptions about the model such as knowing HRF function, assuming AR(1) (or another) process for the residuals, linearity, etc. Despite these limitations, GLM approach to fMRI time-series remains a relatively intuitive and highly flexible tool, especially in light of the many sophisticated methods that have been introduced to resolve assumption violations.

Unlike GLM, ICA relies on the intrinsic structure of the data, no assumptions about the form of the HRF or the possible causes of responses are introduced. Therefore, ICA would be more sensitive in detecting task-related changes in fMRI signal than the traditional GLM based analysis, because ICA uses a data-driven approach, and can reduce noise in the final solution by separating artifacts from real fMRI signal. However ICA has its own challenges. Firstly, ICA decomposition is obtained by means of iterative optimization. This stochastic nature of the process induces a degree of run-to-run variability, so results obtained from such an analysis can differ between analysis runs on even the same data [26]. Secondly, the processes of dimensionality reduction and model order selection are somewhat arbitrary. While approaches exist to optimally select the number of independent components for a given dataset according to statistical criteria, it is important to note that there can be no single, "best" dimension or model order for the underlying neurophysiology [62].

There have been some studies comparing performances of GLM and ICA on empirical data. For example, Robinson et al. [52] used an fMRI data based on clinical study involving chin and hand motion tasks. Their study showed that ICA was capable of cleanly separating activation from motion artifacts in ultra-high field fMRI data which contained stimulus-correlated motion. Some activated regions were evident in ICA results but not in GLM results, indicating not only higher true positive

rate detecting activation but also lower false positive rate in the analysis of motion-contaminated data. In general, studies have shown that ICA can be a valuable tool to detect hidden activity in the brain that cannot be found using a model-based analysis like the GLM [33]. While clearly ICA cannot be used to validate a model, it can give useful hints to understand the brain and help to develop new models and study designs which then can be validated using a classic regression analysis. Thus, it is recommended that investigators use both GLM and sICA in future fMRI studies for a more complete understanding of the brain's functional organization.

Chapter 3

Numerical Analysis

3.1 BOLD Data

The real data used in our analysis comes from an open source data platform and is available at *openneuro.org*. The dataset, called BOLD5000, is a large-scale, slow event-related fMRI dataset. In the experiment almost 5,000 unique images were used.

The images where from one of the following three categories: Scenes, images from the COCO dataset [41], and images from the ImageNet dataset [18]. Images were presented for 1 second, with 9 seconds of fixation cross between trials. Participants were asked to judge whether they liked, disliked, or were neutral about the image. There were four participants in the study. The data available for each participant was obtained through 15 task-related sessions and one session to obtain high resolution anatomical images. The functional images in the dataset were collected using a T2*-weighted gradient. Further details about the scanning parameters are available in the BOLD5000, a public fMRI dataset while viewing 5000 visual images paper [11]. We only investigated one participant to remain in the native space. The focus of our analysis is the contrast between two tasks: fixation versus image viewing (all types of images mixed, unless stated otherwise). We analyzed the data from the first scanning session that contained ten runs for the first subject.

The BOLD5000 dataset used the pipeline developed by the Poldrack lab at Stanford University [11].

3.2 GLM Analysis

3.2.1 Determining HRF

The expected BOLD response can be created using a task indicator function as a convolution with the hemodynamic response function of our choice. The *fmri.stimulus()* function allows us to choose between the canonical, simple, boxcar, and a user defined functions [51]. Here, we provide an empirical comparison of the estimated BOLD response in Figure 3.1. We can observe that the expected BOLD response created using the canonical versus gamma HRF are similar in shape although they differ in the initial values and amplitudes. The BOLD signal measured for the particular voxel has a much greater values than the expected BOLD plots, however, the latter present the expected percent signal change rather than the raw values themselves. The boxcar function is usually used for block design experiments. As the BOLD5000 experiment is a slow event-related experiment, this is not the best choice for modeling purposes. For the analysis in this thesis, we chose to set the function as canonical as it is the default setting and the study of HRF is not the focus of the thesis.

3.2.2 First Level: One Run

Our analysis first focused on the first run of data. We defined image viewing (all image types) versus fixation cross as a task contrast. Images were presented a total of 37 times during this test, which should give the analysis sufficient statistical power to find

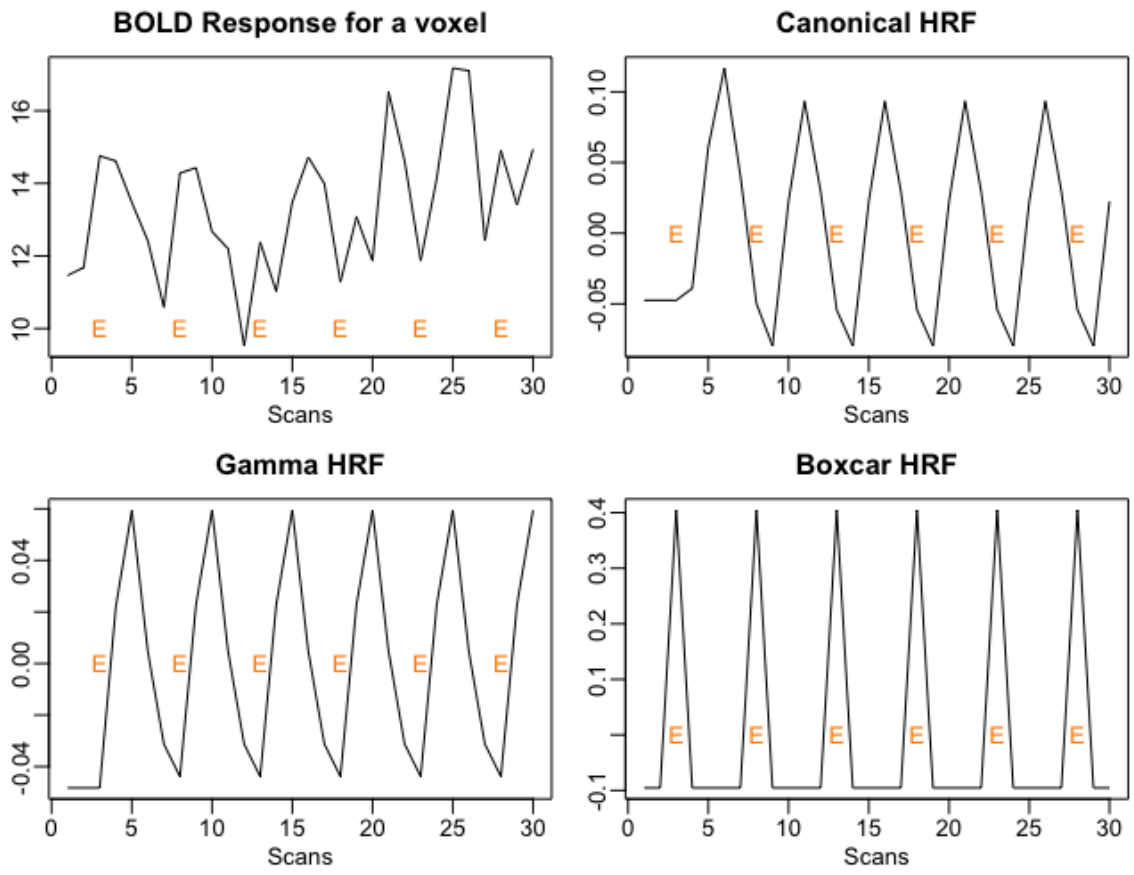

Figure 3.1: Comparison of the BOLD signal for a voxel (top left) and the expected BOLD responses using a Canonical HRF (top right), Gamma HRF (bottom left), and a Boxcar HRF (bottom right).

activated voxels. The *fmri.pvalue()* function allows us to choose the kind of p-value calculation we prefer by changing the mode parameter. It allows for a Bonferroni, False Discovery Rate (FDR), and voxelwise p-value calculation method. Bonferroni adjusts the significance level for multiple testing and is the most conservative option of the three. A fourth option for this parameter is "basic", which uses the estimated resel counts achieved by adaptive smoothing to calculate the p-values. Another parameter of interest is the parameter *alpha*, which allows one to set an α level for tests. We set α to be the default, which is 0.05 [51].

The activation map for the contrast, image viewing versus fixation cross, obtained using the Bonferroni method is showed in Figures 3.2 (a) and 3.3 (a). Like for the FDR map, we observe activation mainly, and almost exclusively in this case, in the occipital lobe, which is consistent with the literature. We can observe that there are fewer voxels that are active using this method. This makes sense as the Bonferroni method is more conservative.

The activation map for the contrast obtained using the FDR method is showed in Figures 3.2 (b) and 3.3 (b). We can observe active voxels mainly located in the occipital lobe, which is consistent with the extensive literature that studied visual tasks. The detailed and commented code is available in the Appendix.

The activation map for the contrast obtained using the voxelwise method is showed in Figures 3.2 (c) and 3.3 (c). We can see activated voxels all over the brain, and even some located outside the subjects' brain. This is evidence that the activated areas obtained are mostly noise. The activation map obtained using the method denoted as "basic" (Figures 3.2 (d) and 3.3 (d)) yielded us a map very similar to the Bonferroni one.

For the contrasts Coco-ImageNet, Coco-Scenes, and Scenes-Imagenet, we found

no significant voxels using the FDR and Bonferroni methods. When we used the voxelwise method, we found active voxels located throughout the brain, and even some outside of it, indicating that mostly noise was found.

3.2.3 Second Level: Multiple Runs

In order to increase the statistical power of the experiment, we can perform the second level of the GLM model by combining 10 runs of data. We only performed group GLM for the task, all images versus fixation cross, as other contrasts did not show significant results for any of the runs separately. We used the function *fmri.metaPar()* to estimate a group map for the ten runs combined. This function performs a voxel-by-voxel analysis and fits a configured linear mixed-effects meta-analytic model [51].

To calculate the p-values and determine the activated voxels we used the same function as for the individual runs. The method "basic" returned no active voxels meanwhile the method "FDR" combined with the default alpha-level of 0.05, returned very large active areas, which seemed inconsistent with the results found from the individual runs. As shown in Figures 3.4 and 3.5, setting the alpha-level to 0.0001 when using the"FDR" mode provided us with maps very similar to the "Bonferroni" mode with a default set alpha-level of 0.05. We can observe that the active voxels still live mainly in the occipital lobe as for the results shown in the first level of the analysis which is consistent with the literature of visual tasks. There also appears some activation, to a lesser extent, in the cerebellum. The cerebellum is mostly known for its role in motor functions but might have been found to be activated because the participants were asked to convey how much they liked the image they viewed by pressing on a button [41].

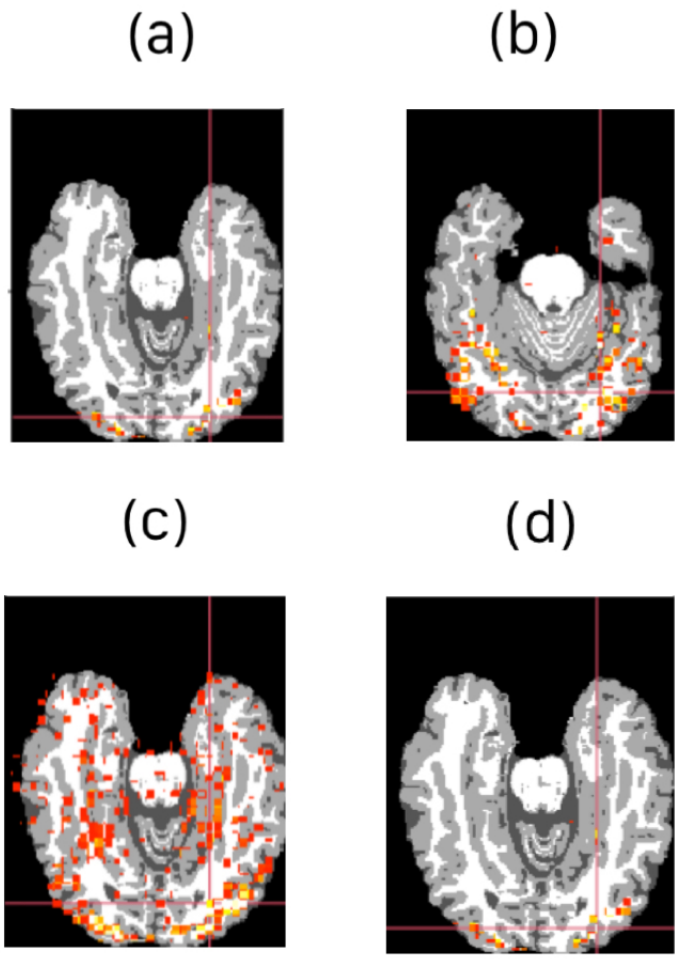

Figure 3.2: Axial view of the map of activated voxels for the first run for the image viewing versus fixation cross task contrast using p-value calculations (a) Bonferroni, (b) FDR, (c) voxelwise, (d) and basic.

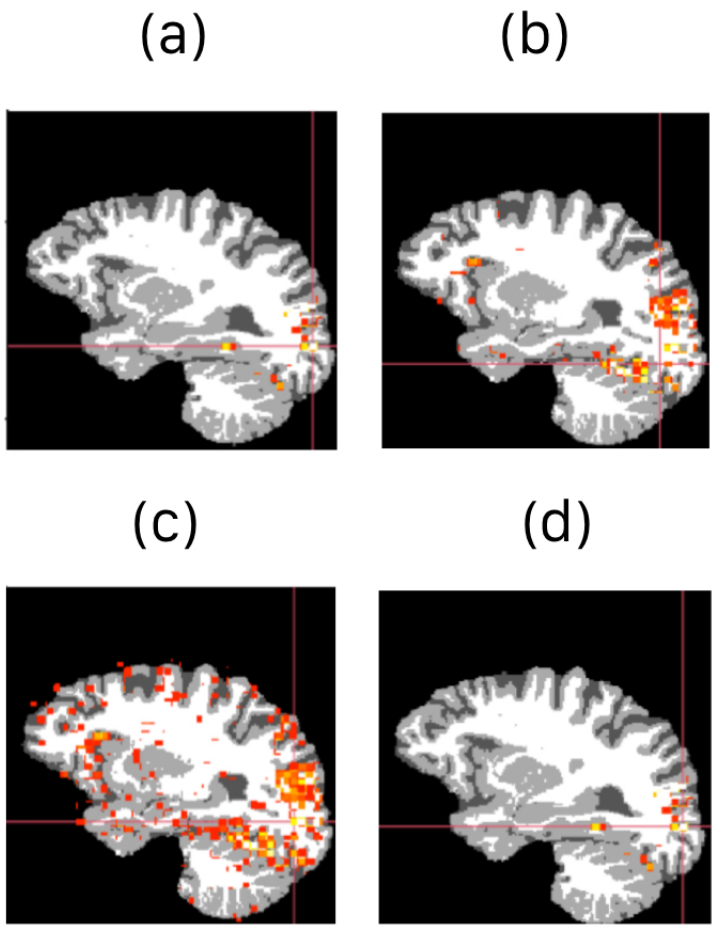

Figure 3.3: Sagittal view of the map of activated voxels for the first run for the image viewing versus fixation cross task contrast using p-value calculations (a) Bonferroni, (b) FDR, (c) voxelwise, (d) and basic.

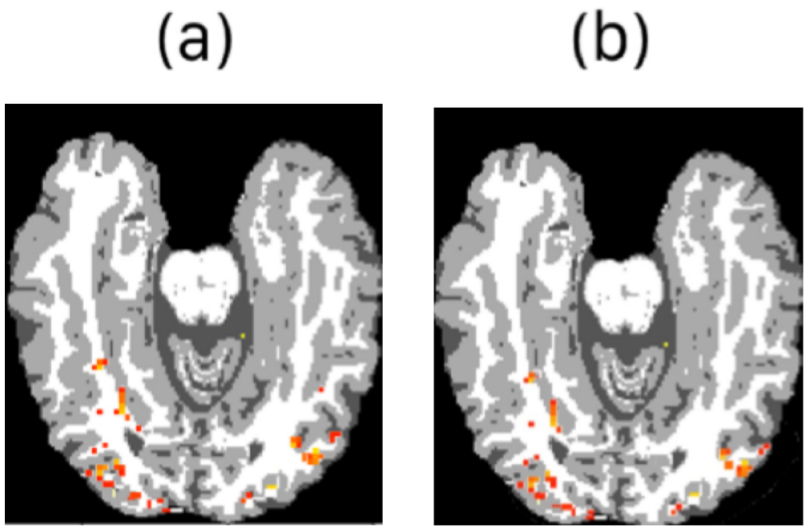

Figure 3.4: Axial view of the map of activated voxels for the first 10 runs combined for the image viewing versus fixation cross task contrast using p-value calculations (a) FDR with alpha set to 0.0001, (b) Bonferroni with default alpha of 0.05.

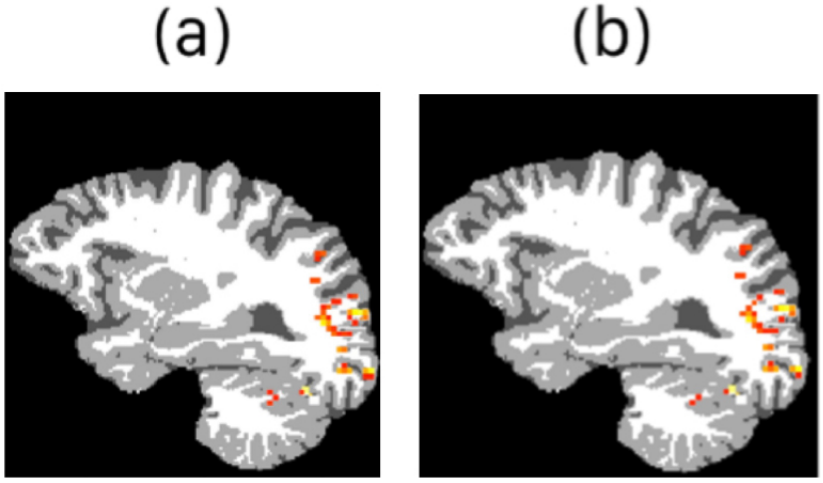

Figure 3.5: Sagittal view of the map of activated voxels for the first 10 runs combined for the image viewing versus fixation cross task contrast using p-value calculations (a) FDR with alpha set to 0.0001, (b) Bonferroni with default alpha of 0.05.

3.3 IC Analysis

3.3.1 One Run

In order to fit ICA to our data, we used the function *fmri.sICA()*, which is a spatial ICA as described in the second chapter. We use "*logcosh*" and "*parallel*" with fastICA in order to accelerate the time elapsed. We also set the parameter smooth to "*TRUE*", so that the resulting residual series are spatially smoothed using a Gaussian kernel with a specified bandwidth, which was set to 4. Here the unit of bandwidth is "Full Width Half Maximum (FWHM)". We set the number of components to estimate to be 20.

In order to distinguish between components of interest that may have a neurological interpretation and nuisance components that describe artifacts from motion, physiological effects or scanner inhomogeneities, etc., IC fingerprints have been introduced and used for automatic IC classification in Martino et al. [19]. Here, independent components are characterized by numerical features of spatial IC and its corresponding time course from the mixing matrix. The characteristics used are kurtosis, skewness, entropy and a clustering index obtained from the spatial part and entropy, first-order autocorrelation, and proportions of five frequency bands in the spectrum of the time component. The characteristics are then normalized, over all components, to be within the unit interval [0, 1].

The star-plot displayed in Figure 3.6 compares the first three components using *fmri.ICAfingerprint()*. This function executes ICA fingerprinting as described by de Martino, with some modifications, such as normalization of values [19, 51]. We can see that the first independent component has the highest spatial entropy (sen-

tropy), degree of clustering in the anatomical space (dclust), and power in the band 2 (power2). The second component has slightly higher temporal entropy (tentropy) than the first one, but both scores remain quite high. These features are specially helpful when doing exploratory data analysis of fMRI data. The degree of clustering score is of particular interest meaningful processes usually have a spatial structure that is well-defined [19]. We thus believe that the first component is the most likely to be associated with image viewing and to have found a higher signal to noise ratio.

The assessment of the different independent components that were estimated indicated that the first component is more likely to be associated with the image viewing task meanwhile other components are more likely to have captured noise.

Figure 3.6 displays a map of the activated voxels in the brain for the first run of data and the first component. We can see that most of the activated voxels live in the occipital lobe. There is some minor activation in the cerebellum and frontal lobe which could be either noise or linked to the subjects having to press a button while viewing an image and making a judgement. There also appears to be some significant activation in the temporal lobe, which was not found using the GLM model. This might be noise. The estimated time series appears to be consistent with the times the events were showed (Figure 3.1 shows the BOLD response and HRF functions for comparison).

3.3.2 Multiple Runs

In order to estimate the independent components for all 10 runs of data combined we used the function *fmri.sgroupICA()*. This function employs a hierarchical clustering algorithm on the combined set of spatial independent components obtained from

the individual time series for each of the 10 runs. The correlations of the independent components are used to calculate a distance matrix [51]. We set the parameter "*thresh*", the threshold for cluster aggregation, to be 0.75. The parameter "*minsize*", minimal size of cluster to considered in IC aggregation, to be 3.

The estimated maps is summarized in Figure 3.8 which shows coronal, saggital, and axial views. We can see that no activation was found in the occipital lobe. Most of the active voxels lived in the temporal lobe and the cerebellum. This is not great evidence that the first independent component found for the group ICA is associated with image viewing. Unfortunately the fingerprint function does not run for group ICA data so we do not have supplementary scores for this component. After exploring the components 2 through 20, we believe that those capture mostly noise.

3.3.3 GLM versus ICA

Our results for the first run of data using GLM (contrast all images versus fixation cross) is consistent with the results displayed for the first IC estimated using ICA. We observe activation mostly in the occipital lobe, with some minor activation in the cerebellum. Additionally the first IC also found activation in the temporal lobe along with some minor activation in the frontal lobe.

On the other hand, the results for the 10 runs of data combined differ more significantly when we compare both methods. The GLM model found activation that was consistent with the results from the first run. ICA, however, found activation located mostly in the cerebellum and temporal lobes and no activation in the occipital lobe, which is what we expected.

Thus, we believe that overall the GLM method results are more consistent with our task of interest (image viewing) than the group ICA for this data application.

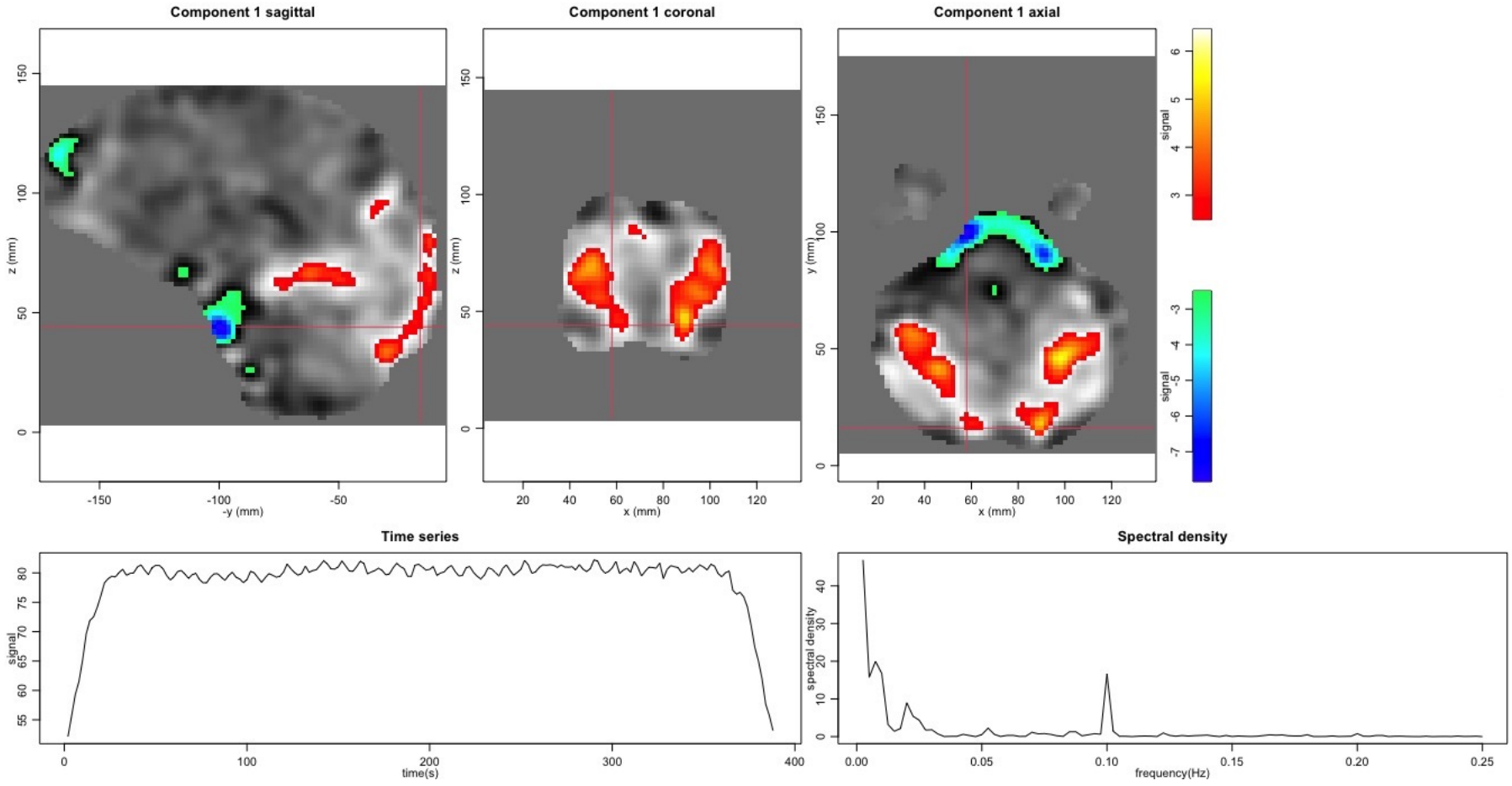

Figure 3.6: Results of the first ICA component for the first run. This plot includes a saggital, coronal, and axial view of the map of activated areas along with the ICA fingerprint and estimated time series with spectral density associated with the component.

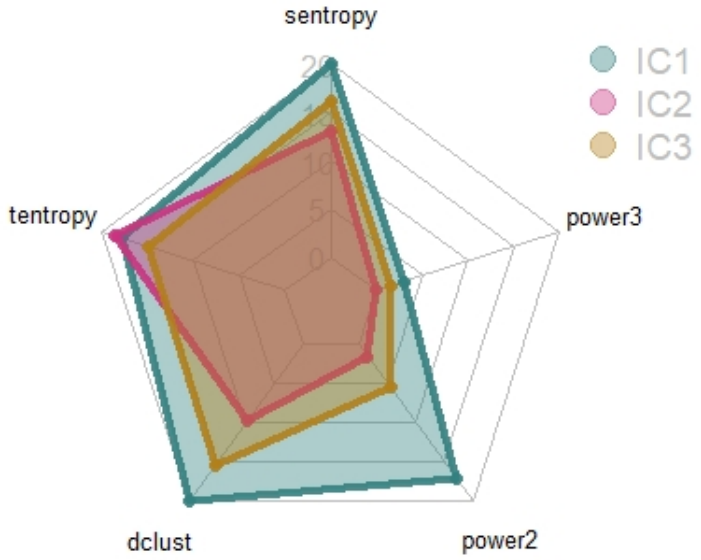

Figure 3.7: Starplot of the ICA fingerprint for run 1 task all images versus fixation cross for components 1 through 3.

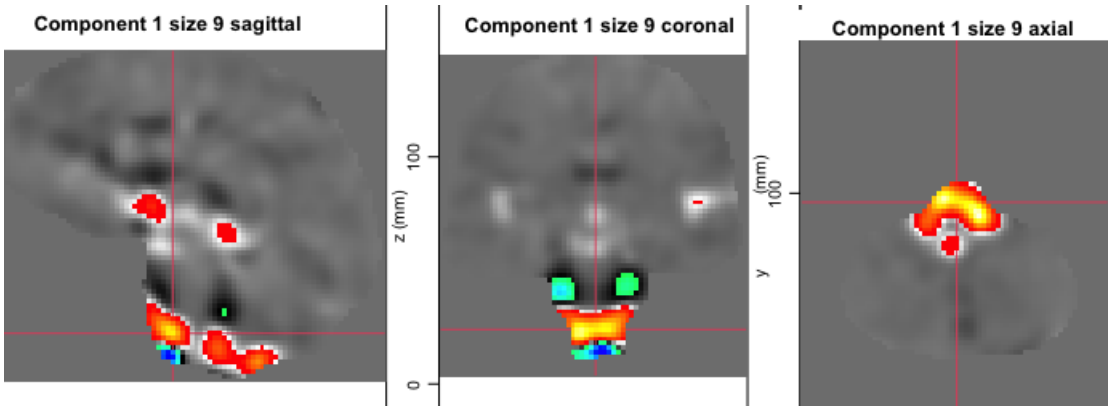

Figure 3.8: Map of activated voxels of the ICA analysis combining 10 runs.

Chapter 4

Conclusion

Functional magnetic resonance imaging is one of the most widely used tools to study the neural underpinnings of human cognition. Applications of statistical methods on fMRI data have provided us with a better understanding of the neural basis of cognitions, emotions, behaviors, and neurologic disorders.

We have found that both the GLM and ICA for the first run of data found significant activation located in the occipital lobe which is consistent with the literature on visual tasks. Both analyses also found activation in the cerebellum which might be related to the experimental design, and more precisely to the motion that participants made while pressing the button. Meanwhile the GLM results for the ten runs of data combined were consistent with the results for the individual runs, the group ICA results were not. The first method found significant activation in the occipital lobe meanwhile the second approach did not, finding activation located mainly in the temporal lobe. It is important to note that it is hard to assess whether the first independent component estimated by the *fmri.sgroupICA()* function is related to the task or not as there is not an implementation of the *ICAfingerprint()* function available for that output.

Given that the interesting statistical questions deepen the more we learn, there are still numerous computational and conceptual challenges that have yet to be fully explored for building more realistic models. Therefore, methodology for statistical

prediction based on fMRI data represents an important area for future research, and preliminary work in this area provides a promising outlook for the potential utility of fMRI data for understanding brain function and its relationship to behavior.

References

[1] Anderson, A., Dinov, I. D., Sherin, J. E., Quintana, J., Yuille, A. L., & Cohen, M. S. (2010). Classification of spatially unaligned fMRI scans. neuroimage, 49(3), 2509-2519.

[2] Arribas, J. I., Calhoun, V. D., & Adali, T. (2010). Automatic Bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from FMRI data. IEEE Transactions on Biomedical Engineering, 57(12), 2850-2860.

[3] Ashburner, J., Barnes, G., Chen, C. C., Daunizeau, J., Flandin, G., Friston, K., ... & Klaas, S. (1994). Statistical parametric mapping.

[4] Bach, F. R., & Jordan, M. I. (2002). Kernel independent component analysis. Journal of machine learning research, 3(Jul), 1-48.

[5] Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. Neural computation, 7(6), 1129-1159.

[6] Bennett, C. M., & Miller, M. B. (2013). fMRI reliability: influences of task and experimental design. Cognitive, Affective, & Behavioral Neuroscience, 13(4), 690-702.

[7] Boscolo, R.H.,Pan, H., & Roychowdhury, V.P. (2001). Non-parametric ICA. San Diego, California

[8] Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. Neuroreport, 9(16), 3735-3739.

[9] Calhoun,V.D., Adali,T., Pearlson, G. & Pekar J.J.(2003). A Method for Testing Conjunctive and Subtractive Hypotheses on Group fMRI Data Using Independent Component Analysis Proc.ISMRM.

[10] Cardoso, J. F., & Souloumiac, A. (1993, December). Blind beamforming for non-Gaussian signals. In IEE proceedings F (radar and signal processing) (Vol. 140, No. 6, pp. 362-370). IET Digital Library.

[11] Chang, N., Pyles, J. A., Marcus, A., Gupta, A., Tarr, M. J., & Aminoff, E. M. (2019). BOLD5000, a public fMRI dataset while viewing 5000 visual images. Scientific data, 6(1), 1-18.

[12] Churchill, N. W., Oder, A., Abdi, H., Tam, F., Lee, W., Thomas, C., ... & Strother, S. C. (2012). Optimizing preprocessing and analysis pipelines for single-subject fMRI. I. Standard temporal motion and physiological noise correction methods. Human brain mapping, 33(3), 609-627.

[13] Churchill, N. W., Spring, R., Afshin-Pour, B., Dong, F., & Strother, S. C. (2015). An automated, adaptive framework for optimizing preprocessing pipelines in task-based functional MRI. PloS one, 10(7), e0131520.

[14] Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI)"brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage, 19(2), 261-270.

[15] Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical research, 29(3), 162-173.

[16] Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. Human brain mapping, 5(5), 329-340.

[17] Demirci, O., Clark, V. P., & Calhoun, V. D. (2008). A projection pursuit algorithm to classify individuals using fMRI data: Application to schizophrenia. Neuroimage, 39(4), 1774-1782.

[18] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.

[19] De Martino, F., Gentile, F., Esposito, F., Balsi, M., Di Salle, F., Goebel, R., & Formisano, E. (2007). Classification of fMRI independent components using IC-fingerprints and support vector machine classifiers. Neuroimage, 34(1), 177-194.

[20] D'Esposito, M., Zarahn, E., & Aguirre, G. K. (1999). Event-related functional MRI: implications for cognitive psychology. Psychological bulletin, 125(1), 155.

[21] Esposito, F., Scarabino, T., Hyvarinen, A., Himberg, J., Formisano, E., Comani, S., ... & Di Salle, F. (2005). Independent component analysis of fMRI group studies by self-organizing clustering. Neuroimage, 25(1), 193-205.

[22] Friston, K. J., Jezzard, P., & Turner, R. (1994). Analysis of functional MRI time-series. Human brain mapping, 1(2), 153-171.

[23] Fan, Y., Shen, D., & Davatzikos, C. (2006, June). Detecting cognitive states from fMRI images by machine learning and multivariate classification. In 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06) (pp. 89-89). IEEE.

[24] Friston, K. J., Zarahn, E. O. R. N. A., Josephs, O., Henson, R. N., & Dale, A. M. (1999). Stochastic designs in event-related fMRI. Neuroimage, 10(5), 607-619.

[25] Friston, K.J., Ashburner, J., Kiebel, S.J., Nichols, T.E., & Penny, W.D. (2007). Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press, New York.

[26] Himberg, J., Hyvärinen, A., & Esposito, F. (2004). Validating the independent components of neuroimaging time series via clustering and visualization. Neuroimage, 22(3), 1214-1222.

[27] Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. Biometrika, 75(4), 800-802.

[28] Huettel, S., Song, A., & McCarthy, G. (2014). Functional Magnetic Resonance Imaging, 3rd edn. Sinauer Associates, Inc.

[29] Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE transactions on Neural Networks, 10(3), 626-634.

[30] Hyvarinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. Neural computation, 9(7), 1483-1492.

[31] Hyvarinen, A., Karhunen, J., & Oja, E. (2001). Independent component analysis and blind source separation.

[32] Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. Neuroimage, 62(2), 782-790.

[33] Keck, I. R., Theis, F. J., Gruber, P., Lang, E. W., Specht, K., & Puntonet, C. G. (2004, July). 3D spatial analysis of fMRI data: a comparison of ICA and GLM analysis on a word perception task. In 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541) (Vol. 3, pp. 2495-2499). IEEE.

[34] Lazar, N. (2008). The statistical analysis of functional MRI data. Springer Science & Business Media.

[35] Lazar, N. A., Eddy, W. F., Genovese, C. R., & Welling, J. (2001). Statistical issues in fMRI for brain imaging. International Statistical Review, 69(1), 105-127.

[36] Lazar, N.A. (2017). Corrections for multiplicity in functional neuroimaging data. In Handbook of Neuroimaging Data Analysis. CRC Press.

[37] Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). Review of methods for functional brain connectivity detection using fMRI. Computerized medical imaging and graphics, 33(2), 131-139.

[38] Lindquist, M. A. (2008). The statistical analysis of fMRI data. Statistical science, 23(4), 439-464.

[39] Liu, T. T., & Frank, L. R. (2004). Efficiency, power, and entropy in event-related FMRI with multiple trial types: Part I: theory. NeuroImage, 21(1), 387-400.

[40] Martínez-Ramón, M., Koltchinskii, V., Heileman, G. L., & Posse, S. (2006). fMRI pattern classification using neuroanatomically constrained boosting. Neuroimage, 31(3), 1129-1141.

[41] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755). Springer, Cham.

[42] McKeown, M. J., Jung, T. P., Makeig, S., Brown, G., Kindermann, S. S., Lee, T. W., & Sejnowski, T. J. (1998). Spatially independent activity patterns in functional MRI data during the Stroop color-naming task. Proceedings of the National Academy of Sciences, 95(3), 803-810.

[43] McKeown, M. J., Makeig, S., Brown, G. G., Jung, T. P., Kindermann, S. S., Bell, A. J., & Sejnowski, T. J. (1998). Analysis of fMRI data by blind separation into independent spatial components. Human brain mapping, 6(3), 160-188.

[44] Meltzer, J. A., Zaveri, H. P., Goncharova, I. I., Distasio, M. M., Papademetris, X., Spencer, S. S., ... & Constable, R. T. (2008). Effects of working memory load on oscillatory power in human intracranial EEG. Cerebral Cortex, 18(8), 1843-1855.

[45] Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage, 11(6), 735-759.

[46] Mumford, J. A., & Nichols, T. (2009). Simple group fMRI modeling and inference. Neuroimage, 47(4), 1469-1475.

[47] Parker, D. B., & Razlighi, Q. R. (2019). The benefit of slice timing correction in common fMRI preprocessing pipelines. Frontiers in neuroscience, 13, 821.

[48] Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (Eds.). (2011). Statistical parametric mapping: the analysis of functional brain images. Elsevier.

[49] Poggio, G. F. (1972). Spatial properties of neurons in striate cortex of unanesthetized macaque monkey. Investigative Ophthalmology, 11(5), 368-377.

[50] Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). Handbook of functional MRI data analysis. Cambridge University Press.

[51] Polzehl, J., Tabelow, K. (2007). fmri: A Package for Analyzing fmri Data.

[52] Robinson, S. D., Schöpf, V., Cardoso, P., Geissler, A., Fischmeister, F. P. S., Wurnig, M., ... & Beisteiner, R. (2013). Applying independent component analysis to clinical FMRI at 7 T. Frontiers in human neuroscience, 7, 496.

[53] Sarty, G. E. (2007). Computing brain activity maps from fMRI time-series images.

[54] Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage, 23, S208-S219.

[55] Strother, S. C., & Churchill, N. (2017). Neuroimage preprocessing. Handbook of neuroimaging data analysis, 264-308.

[56] Team, R. C. (2017). R Foundation for Statistical Computing; Vienna, Austria: 2016. R: A language and environment for statistical computing.

[57] Wang, X., Hutchinson, R., & Mitchell, T. M. (2003). Training fMRI classifiers to detect cognitive states across multiple human subjects. NIPS03, 16.

[58] Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., & Evans, A. C. (2002). A general statistical analysis for fMRI data. Neuroimage, 15(1), 1-15.

[59] Worsley, K. J. (2005). Spatial smoothing of autocorrelations to control the degrees of freedom in fMRI analysis. NeuroImage, 26(2), 635-641.

[60] Zhang, H., Luo, W. L., & Nichols, T. E. (2006). Diagnosis of single-subject and group fMRI data with SPMd. Human brain mapping, 27(5), 442-451.

[61] Zhang, J., Anderson, J. R., Liang, L., Pulapura, S. K., Gatewood, L., Rottenberg, D. A., & Strother, S. C. (2009). Evaluation and optimization of fMRI single-subject processing pipelines with NPAIRS and second-level CVA. Magnetic resonance imaging, 27(2), 264-278.

[62] Zuo, X. N., Kelly, C., Adelstein, J. S., Klein, D. F., Castellanos, F. X., & Milham, M. P. (2010). Reliable intrinsic connectivity networks: test–retest evaluation using ICA and dual regression approach. Neuroimage, 49(3), 2163-2177.

Appendix A

R Code

```
#Load in necessary libraries


#If you use MacOS make sure you download R from


#https://cran.r-project.org/bin/macosx/ directly



#so tcltk is included in the installation
rm(list=ls())
library(fastICA)
library(fmri)
library(oro.nifti)
library(Rcmdr)
library(tkrplot)


#base package no need to install
library(tcltk)


#download the data from
```

```
#https://openneuro.org/datasets/ds001499/versions/1.3.0

#We recommend having a file structure as follows:
     #Sub-CSI1/ folder for CSI1 data
      #/anat/ where your anatomical image is located
      #/func/ where your functional data is located
        #/runi i = 1,...,10




###Load the Anatomic data
#set your working directory to where your Sub-CSI1

data is located

setwd("<PATH_TO_DATA>/Sub-CSI1/")

ana= readNIfTI("anat/derivatives_fmriprep_sub-CSI1_anat_sub-CSI1_T1w_
   ↪ dtissue.nii")



Vbold=Cbold <- array(0, dim = c(71, 89, 72, 10))
```

```
#set working directory to your func folder
setwd("<PATH_TO_DATA>/Sub-CSI1/func")
runs=c("s1r1","s1r2","s1r3","s1r4","s1r5","s1r6","s1r7","s1r8","s1r9",
    ↪ "s1r10")

##Task 1: all images versus fixation cross
setwd("<PATH_TO_DATA>/Sub-CSI1/func/")
fm<-"/derivatives_fmriprep_sub-CSI1_ses-01_func_sub-CSI1_ses-01_task
    ↪ -5000scenes_run-01_bold_space-T1w_brainmask.nii.gz"
mask1 <-readNIfTI(fm)>0



i=10
rn=runs[i]
setwd(rn)
getwd()



###need to modify following file reading for runs where i!=1
fr <-"derivatives_fmriprep_sub-CSI1_ses-01_func_sub-CSI1_ses-01_task
    ↪ -5000scenes_run-10_bold_space-T1w_preproc.nii.gz"
```

```
sub1_s1_ri <- readNIfTI(fr, reorient = FALSE)
s1r1_events = read.table("sub-CSI1_ses-01_func_sub-CSI1_ses-01_task
    ↪ -5000scenes_run-10_events.tsv", sep = '\t', header = TRUE)

onsets_s1r1 = s1r1_events$onset
onsets_s1r1_scans =round(onsets_s1r1/2, digits = 0)
all_images = fmri.stimulus(scans = 194,onsets = onsets_s1r1_scans,
    ↪ durations=c(rep(0.5,37)), TR=2, type="canonical")
x_all = fmri.design(all_images)

ds <- oro2fmri(sub1_s1_ri)
ds$mask <- mask1

spm_all = fmri.lm(ds,mask=ds$mask,x_all,actype = "smooth")



Cbold[,,,i] <- spm_all$cbeta
Vbold[,,,i] <- spm_all$var

###In fmri.lm, If actype
#%in% c("ac","accalc","smooth") an AR(1) model is fitted, in each
    ↪ voxel, to the time series
#of residuals. The estimated AR-coefficients are corrected for bias.
    ↪  If actype=="smooth" the estimated
```

```
#AR-coefficients are spatially smoothed. If actype %in% c("ac","
   ↪ smooth") the linear model
#is pre-withened using the estimated (and possibly smoothed) AR-
   ↪ coefficients. Parameter and variance
#estimates are then obtained from the pre-whitened data.

pvalue <- fmri.pvalue(spm_all, mode="FDR")

#The parameter mode allows for different kinds of p-value
   ↪ calculation. mode="voxelwise" refers to
#voxelwise tests while mode="Bonferroni" adjusts the significance
   ↪ level for multiple testing.

#alternative is mode="FDR" specifying signal detection by False
   ↪ Discovery Rate (FDR) with proportion of false positives level
   ↪ specified by alpha. The other choices apply results on
   ↪ excursion

#sets of random fields (Worsley 1994, Adler 2003) for smoothed SPM?s
   ↪ . "basic" corresponds to a
#global definition of the resel counts based on the amount of
   ↪ smoothness achieved by an equivalent
#Gaussian filter.
```

```
plot.new()
plot(pvalue,view="orthographic",template=ana,mask=mask1) #all images
    ↪ ###RUN1


save(spm_all,file=paste(rn,"Rdata",sep="."))
save("Cbold","Vbold",file="VC_all.Rdata")


###separate images
setwd("<PATH_TO_DATA>/Sub-CSI1/func")
ttt <- s1r1_events
ntrials <- dim(ttt)[1]


indImagenet <- (1:ntrials)[(ttt$ImgType=="imagenet") | (ttt$ImgType=="
    ↪ rep_imagenet")]


indCoco <- (1:ntrials)[(ttt$ImgType=="coco")| (ttt$ImgType=="rep_coco"
    ↪ ) ]


indScenes <- (1:ntrials)[(ttt$ImgType=="scenes")]
onsets <- ttt$onset
duration <- ttt$duration
```

```
HRFImagenet <- fmri.stimulus(scans=194, onsets=onsets[indImagenet],
    ↪ durations=duration[indImagenet],TR =2, times=TRUE)

HRFCoco <- fmri.stimulus(scans=194,

onsets[indCoco],duration[indCoco],TR =2, times=TRUE)
HRFScenes <- fmri.stimulus(scans=194,

onsets[indScenes],duration[indScenes],TR =2, times=TRUE)

HRF <- cbind(HRFImagenet,HRFCoco,HRFScenes)

###Imagenet versus Coco
x_sep = fmri.design(HRF)
spm_sep = fmri.lm(ds,mask=ds$mask,x_sep,contrast=c(1,-1,0,0,0,0),
    ↪ actype = "smooth")
pvalue_sep <- fmri.pvalue(spm_sep, mode="FDR") ###change mode and see
    ↪  if any voxel is defined
plot(pvalue_sep)
save(spm_sep,file=paste(paste(rn,"sepImCo",sep=""),"Rdata",sep="."))



plot.new()
```

```
plot(pvalue,view="orthographic",template=ana,mask=mask1) #all images
    ↪ ###RUN1



####Combining all runs for the all_events ---later!!!
#dim(spm_all$cbeta) :71,89,72
#dim(spm_all_r2$cbeta) :71,89,72



setwd("<PATH_TO_DATA>/Sub-CSI1/func")
runs=c("s1r1","s1r2","s1r3","s1r4","s1r5","s1r6","s1r7","s1r8","s1r9",
    ↪ "s1r10")


for (i in 1:10){
  print(i)
  setwd("<PATH_TO_DATA>0/Sub-CSI1/func")
  rn=runs[i]
  setwd(rn)
  getwd()
  load(list.files(pattern="RData")[1])


  Cbold[,,,i] <- spm_all$cbeta
  Vbold[,,,i] <- spm_all$var
```

```
  rm("spm_all")}


Vbold[Vbold==0]=1e-04


setwd("<PATH_TO_DATA>")
save(Cbold,Vbold,file="GLM_sub1_allruns.RData")



####2nd stage with metaPar
rm(list=ls())
library(fmri)
load("<PATH_TO_DATA>/GLM_sub1_allruns.RData")
out=fmri.metaPar(Cbold,Vbold,method = "FE", cluster = 1,knha = FALSE)



pval= fmri.pvalue(out,na.rm=TRUE,mode="FDR")


plot.fmripvalue(pval)




######## ICA
```

```
rm(list=ls())

###Anatomic data

runs=c("s1r1","s1r2","s1r3","s1r4","s1r5","s1r6","s1r7","s1r8","s1r9",
    ↪ "s1r10")



###set working directory

setwd("<PATH_TO_DATA>/Sub-CSI1/func")


fm<-"derivatives_fmriprep_sub-CSI1_ses-01_func_sub-CSI1_ses-01_task
    ↪ -5000scenes_run-01_bold_space-T1w_brainmask.nii.gz"

mask1 <-readNIfTI(fm)>0


i=1

rn=runs[i]

setwd(rn)

getwd()


#load in preprocessed data for run1

sub1_s1_ri= read.NIFTI("sub-CSI1_ses-01_task-5000scenes_run-01_bold_
    ↪ space-T1w_preproc") #run


ICA_sub1r1= fmri.sICA(sub1_s1_ri,mask1, ncomp= 20, degree = 3,bws = 8,
    ↪  bwt = 4, unit = "FWHM") #r1
```

```
ICA_sub1_f1= ICAfingerprint(ICA_sub1r1,plot=TRUE) #f1 and r1
plot.new()
#make sure your plot window is large enough
plot(ICA_sub1_f1,thresh=3)
#save(ICA_sub1r1,file="s1r1ica_new.RData") #r1 f1 and r1


####Plotting IC-fingerprints:


# Library
library(fmsb)


# Create data:


data <- as.data.frame(ICA_sub1_f1$fingerprint)
data<-data[1:3,c(3,6,4,8,9)]
data <- rbind(rep(max(data),5) , rep(min(data),5) , data)


rownames(data)=c("r1","r2",paste("IC" , c(1:(nrow(data)-2)) , sep=""))


# Color vector
colors_border=c( rgb(0.2,0.5,0.5,0.9), rgb(0.8,0.2,0.5,0.9) , rgb
   ↪ (0.7,0.5,0.1,0.9) )
colors_in=c( rgb(0.2,0.5,0.5,0.4), rgb(0.8,0.2,0.5,0.4) , rgb
   ↪ (0.7,0.5,0.1,0.4) )
```

```
par(mar=c(1,1,1,1))
# plot with default options:
radarchart( data , axistype=1 ,
            #custom polygon
            pcol=colors_border , pfcol=colors_in , plwd=4 , plty=1,
            #custom the grid
            cglcol="grey", cglty=1, axislabcol="grey", caxislabels=seq
               ↪ (0,20,5), cglwd=0.8,
            #custom labels
            vlcex=0.8
)

# Add a legend
legend(x=1, y=1.2, legend = rownames(data[-c(1,2),]), bty = "n", pch
   ↪ =20 , col=colors_in , text.col = "grey", cex=1.2, pt.cex=3)



####Group ICA: combining the results



rm(list=ls())
library(fmri)
```

```
#Load in your data for all 10 runs
setwd("<PATH_TO_DATA>/Sub-CSI1/func")



icaobjlist=list(ICA_sub1r1,ICA_sub1r2,ICA_sub1r3,

ICA_sub1r4,ICA_sub1r5,ICA_sub1r6,ICA_sub1r7,ICA_sub1r8,ICA_sub1r9,

ICA_sub1r10)



ICA_all=fmri.sgroupICA(icaobjlist, thresh =0.75, minsize=3)
#generate a plot of the first component
plot(ICA_all,comp=1,thresh=3)
```